Synergistic multi-spectral CT reconstruction with directional total variation

Cueva, Evelyn

2021-08-23

Cueva , E , Meaney , A , Siltanen , S & Ehrhardt , M J 2021 , ' Synergistic multi-spectral CT reconstruction with directional total variation ' , Philosophical transactions - Royal Society. Mathematical, Physical and engineering sciences , vol. 379 , no. 2204 , 20200198 . https://doi.org/10.1098/rsta.2020

http://hdl.handle.net/10138/337895
https://doi.org/10.1098/rsta.2020

SYNERGISTIC MULTI-SPECTRAL CT RECONSTRUCTION WITH DIRECTIONAL TOTAL VARIATION

EVELYN CUEVA*, ALEXANDER MEANEY†, SAMULI SILTANEN†, AND MATTHIAS J. EHRHARDT‡

Abstract. This work considers synergistic multi-spectral CT reconstruction where information from all available energy channels is combined to improve the reconstruction of each individual channel, we propose to fuse this available data (represented by a single sinogram) to obtain a polyenergetic image which keeps structural information shared by the energy channels with increased signal-to-noise-ratio. This new image is used as prior information during a channel-by-channel minimization process through the directional total variation. We analyze the use of directional total variation within variational regularization and iterative regularization. Our numerical results on simulated and experimental data show improvements in terms of image quality and in computational speed.

Key words. undersampled data, multi–energy CT, directional total variation, linearized Bregman iteration, high-resolution reconstruction.

1. Introduction.

1.1. Undersampling in multi-spectral CT. Computed tomography (CT) is a widely used technique in many different fields of science and industry; for example in medicine, it enables visualizing the internal structure of a patient. The principle of this technique is to study the attenuation of X-rays when they pass through the target object [41, 8]. Despite the potential usefulness of CT, the X-ray source produces a single energy spectrum and the detector does not discriminate between photon energies. As a consequence, two tissues whose elemental composition are different might be indistinguishable in the resulting CT image [52, 53, 39]. The latter makes it difficult to identify and classify different tissues and motivates the multi-spectral techniques based on new scanner technologies [36, 2, 37].

Dual and multi-spectral CT use different technical approaches for acquiring multi-energetic data, *e.g.*, the rapid tube potential switching, multilayer detectors, or dual (multi) X-ray sources [62, 59, 33, 30]. This multi-energetic data provides much more information about the tissue composition allowing to differentiate its constituent materials [39] but, in addition, the measurement process needs a balance between radiation dose, acquisition time and image quality. A reduction in radiation dose is achieved by reducing the number of views in the acquisition which, in turn, decreases the spatial resolution [38, 24, 35]. A recent study proposes to reconstruct a multi-spectral CT image by reducing the dose in each energy window, when just a limited and non-overlapping range of angles is observed [55]. As the resolution of the reconstructions is affected by this lack of measurements, small objects cannot be reconstructed and the resulting images are affected by the presence of artifacts [29, 21]. Many advanced reconstruction techniques have been proposed to simultaneously or independently reconstruct an spectral-image in this scenario [22, 32, 42, 58, 55, 28]. For example, variational methods are commonly used since they allow to directly incorporate prior information and constraints into the model [49]. In addition, regularizers can be added as part of the objective functional or in the optimization process, to overcome ill-posedness [48]. The expected structural correlation between different energy levels has motivated the use of structural priors to improve joint or individual reconstructions [31, 44, 32, 26, 55], some of them based on level sets methods [16]. For example, in [32], the authors proposed a joint image reconstruction method for x-ray spectral computed tomography. Their proposed dTV-p method propagates the structural information along channels using direction total variation (dTV) with a reference image chosen by a probability mass function and considering a wide number of channels (70) for a synthetic phantom. Direction total variation has been successfully used in several other medical imaging applications [18, 17, 19] as in hyperspectral remote sensing [7]. Another similar regularization applied in joint reconstructions to include structural dependence between channels, is the total nuclear variation (TVN) presented in [44, 45].

*Research Center on Mathematical Modelling (MODEMAT), Escuela Politécnica Nacional, Quito, Ecuador (ecueva@uchile.cl)

†Department of Mathematics and Statistics, University of Helsinki, Finland

‡Institute for Mathematical Innovation, University of Bath, United Kingdom

1.2. Main contribution. We propose a novel reconstruction technique to solve the undersampling problem in multi-spectral CT, where information from all available energy channels is combined to obtain a polychromatic image. The latter keeps the structural information shared by the energy channels and is used to improve the reconstruction of each individual channel (channel-by-channel reconstruction) using directional total variation (dTV). We explore variational and iterative regularization methods, specifically, the forward-backward splitting algorithm (FBS) [13, 12] and Linearized Bregman iterations [43, 60, 61, 5, 14] to solved the undersampling problem using simulated and experimental data. The combination of these methods and dTV shows improvements in terms of image quality and computational speed compared, for example, to joint reconstruction techniques as TNV.

In section 2, we describe the inverse problem behind multi-spectral CT data seen as a minimization problem. Later, in section 3, we present the variation and iterative regularization of the inverse problem and we describe the FBS algorithm and Bregman iterations to solved them,respectively. We include total variation and directional total variation regularizers needed during the regularization process. The last section is devoted to present numerical results using synthetic and real data. Here, we specify all the settings needed during the reconstruction.

2. Inverse problem. Multi-spectral CT aims to recover energy-dependent attenuation maps $\boldsymbol{u}_k$ of a target object for energies E_k with $k = 1, ..., K$. The acquisition method considers X-ray projections using only a limited set of angles, *i.e.*, we want to reconstruct $\boldsymbol{u}_k \in \mathbb{R}^N$ given data $\boldsymbol{b}_k \in \mathbb{R}^M$ where $M \ll N$. When a considerable amount of measurements is available, classical methods such as filtered back projection, Kaczmarz iterations or iterative techniques can be used to solve an associated linear system of the form $A\boldsymbol{u}_k = \boldsymbol{b}_k$ or the associated least squares problem (see *e.g.* [41, 10])

$$\min_{\boldsymbol{u}_k \in \mathbb{R}^N} \frac{1}{2} \|A\boldsymbol{u}_k - \boldsymbol{b}_k\|_2^2, \tag{2.1}$$

where A is the forward operator (a matrix in the discrete case) that relates the image $\boldsymbol{u}_k$ to the given data $\boldsymbol{b}_k$. The ill-posedness of this inverse problem makes a direct inversion of the matrix A unstable even for a suitable number of measurements. The undersampling scenario is even more challenging, since $M \ll N$, the system is under-determined.

For 2D CT, $M = m_1 \cdot m_2$ where m_1 is the number of angles and m_2 is the number of detectors, and $N = n_1 \cdot n_2$, where n_1 and n_2 are the number of rows and columns of $\boldsymbol{u}_k$ (considered as a matrix of pixels), respectively.

2.1. Forward model. We recall the forward modelling for multi-spectral CT. For a fixed energy channel E_k, an initial intensity $I_i^0(E_k)$ of X-rays is emitted along a line L_i (from source to detector) given a final intensity I_i^1, for $i = 1, \dots M$. The discretized linear model use for reconstructing a vectorized image $\boldsymbol{u}(E_k)$ of N pixels (see, *e.g* [55]) is given by

$$b_{ik} := -\ln\left(\frac{I_i^1}{I_i^0(E_k)}\right) \approx \sum_{j=1}^{N} a_{ij} u_j(E_k). \tag{2.2}$$

In (2.2), $u_j(E_k)$ is the value of $\boldsymbol{u}(E_k)$ in the corresponding pixel j, and a_{ij} is the length of the intersection of the i-th line and the j-th pixel.

Based on the discretization presented above for each energy E_k, we establish the forward model for the projection data

$$\boldsymbol{b}_k = A_k \boldsymbol{u}_k + \boldsymbol{e}_k, \qquad k = 1, \dots, K, \tag{2.3}$$

where $A_k \in \mathbb{R}^{M \times N}$, is a matrix with components a_{ij}, the vector $\boldsymbol{u}_k \in \mathbb{R}^N$ has components u_j, $\boldsymbol{e}_k$ models the measurements noise and, $\boldsymbol{b}_k$ is the vector of measurements b_{ik} in (2.2) for the fixed energy level k. The matrix A_k represents the discretization of the X-ray transform for a particular projection geometry.

From now on, we omit the energy sub-index in (2.3) since we will solve an independent problem for each energy channel.

3. Regularization. In this section we discuss the regularizers used in this work, total variation and directional total variation, and how these can be used to regularize an inverse problem. To this end we consider variational regularization and iterative regularization based on Bregman iterations.

3.1. Regularizers.

3.1.1. Total Variation. The total variation (TV) regularization has been widely studied due to its edge-preserving properties [47]. TV is well-known to promote piecewise constant images with sharp edges.

To define the (discrete) total variation for an image $\boldsymbol{u}$ of $n_1 \times n_2$ pixels, *i.e.*, $\boldsymbol{u} \in \mathbb{R}^N$ with $N = n_1 \cdot n_2$, we first introduced the discrete gradient $\nabla\colon \mathbb{R}^N \to (\mathbb{R}^2)^N$ based on a finite difference scheme acting on the image pixels, as follows

$$((\nabla \boldsymbol{u})_j)_1 = \begin{cases} u_{j+r(j)} - u_j, & \text{if } r(j) < N \\ 0, & \text{otherwise} \end{cases} \qquad ((\nabla \boldsymbol{u})_j)_2 = \begin{cases} u_{j+s(j)} - u_j, & \text{if } s(j) < N \\ 0, & \text{otherwise} \end{cases}$$

where $r(j)$ is the right neighbour of pixel j, and $s(j)$ is the neighbour below pixel j. Then

$$\mathrm{TV}(\boldsymbol{u}) = \sum_{j=1}^{N} \left(\left[((\nabla \boldsymbol{u})_j)_1 \right]^2 + \left[((\nabla \boldsymbol{u})_j)_2 \right]^2 \right)^{1/2}.$$

3.1.2. Directional Total Variation. While TV is a powerful regularizer, it is unclear how additional structural a-priori information can be included. To this end we utilize the directional total variation (dTV) proposed in [17]. Let $\boldsymbol{\xi} \in (\mathbb{R}^2)^N$ be a vector field with $\|\boldsymbol{\xi}_i\| \leq \eta < 1$. We denote by $\mathbf{P} \in \left(\mathbb{R}^{2\times 2}\right)^N$, $\mathbf{P}_i := \mathbf{I} - \boldsymbol{\xi}_i \otimes \boldsymbol{\xi}_i$ an associated matrix-field, where $\mathbf{I}$ is the 2×2 matrix and $\otimes$ represents the outer product of vectors. Then $\mathrm{dTV}\colon \mathbb{R}^N \to \mathbb{R}$ is defined as

$$\mathrm{dTV}(\boldsymbol{u}; \boldsymbol{v}) = \sum_j \|\mathbf{P}_j (\nabla \boldsymbol{u})_j\|, \tag{3.1}$$

where $\mathbf{P}_j$ implicitly depends on $\boldsymbol{v}$ by means of $\boldsymbol{\xi}$.

Some interpretations of dTV are detailed in [18, 7]. We briefly describe some useful properties of this functional using the explicit expression $\mathbf{P}_j(\nabla \boldsymbol{u})_j = (\nabla \boldsymbol{u})_j - \langle \boldsymbol{\xi}_i, (\nabla \boldsymbol{u})_j \rangle \boldsymbol{\xi}_i$. We observe two particular cases:

$$\mathbf{P}_j(\nabla \boldsymbol{u})_j = \begin{cases} (1 - \|\boldsymbol{\xi}_j\|^2)(\nabla \boldsymbol{u})_j, & \text{if } (\nabla \boldsymbol{u})_j \text{ is parallel to } \boldsymbol{\xi}_j \\ (\nabla \boldsymbol{u})_j, & \text{if } (\nabla \boldsymbol{u})_j \text{ is perpendicular to } \boldsymbol{\xi}_j. \end{cases}$$

So, when we minimize $\mathrm{dTV}(\boldsymbol{u})$, we are favouring $\boldsymbol{u}$ such that its gradient is collinear to the direction $\boldsymbol{\xi}_i$ as long as $\|\boldsymbol{\xi}_i\| \neq 0$. We note that a vanishing gradient $\nabla u = 0$ always leads to a smaller function value such that no artificial jumps are enforced.

In order to incorporate dTV into our model, we define the vector field below based on the known image $\boldsymbol{v} \in \mathbb{R}^N$ by

$$\boldsymbol{\xi}_j = \eta \frac{(\nabla \boldsymbol{v})_j}{\|(\nabla \boldsymbol{v})_j\|_\varepsilon} \tag{3.2}$$

with $\|\boldsymbol{u}\|_\varepsilon = \sqrt{\|\boldsymbol{u}\|^2 + \varepsilon^2}$. The parameter $\varepsilon > 0$ avoids singularities when $(\nabla \boldsymbol{v})_j = 0$ and η is an edge parameter related to the size of an edge.

Figure 3.1 shows an example which compared TV and dTV. In contrast to TV, dTV only penalizes edges which are missing in the side information.

3.2. Variational regularization.

A strategy to reconstruct $\boldsymbol{u} := \boldsymbol{u}_k$ in (2.3) is to solve

$$\boldsymbol{u}^* \in \operatorname*{arg\,min}_{\boldsymbol{u} \in \mathbb{R}^N} \left\{ \frac{1}{2} \|A\boldsymbol{u} - \boldsymbol{b}\|_2^2 + \alpha J(\boldsymbol{u}) + \iota_{[0,\infty)^N}(\boldsymbol{u}) \right\}. \tag{3.3}$$

The first term in (3.3) is called the data-fit and forces $A\boldsymbol{u}$ to stay close to the data, and the regularizer J promotes stability of the inversion. The parameter $\alpha > 0$ balances the data-fit term and the regularization provided by J. We will use TV and dTV as J. The additional term $\iota_{[0,\infty)^N}(\boldsymbol{u})$ is included to impose a nonnegative constraint for each component of the solution $\boldsymbol{u}^*$ and is defined as:

$$\iota_{[0,\infty)^N}(\boldsymbol{u}) = \begin{cases} 0, & \text{if } u_j \geq 0 \\ \infty, & \text{otherwise.} \end{cases}$$

Depending on the type of regularization that we choose, we define the following functions:

$$G_{\mathrm{TV}}(\boldsymbol{u}) = \alpha\, \mathrm{TV}(\boldsymbol{u}) + \iota_{[0,\infty)}(\boldsymbol{u}), \tag{3.4}$$

$$G_{\mathrm{dTV}}(\boldsymbol{u}) = \alpha\, \mathrm{dTV}(\boldsymbol{u}, \boldsymbol{v}) + \iota_{[0,\infty)}(\boldsymbol{u}). \tag{3.5}$$

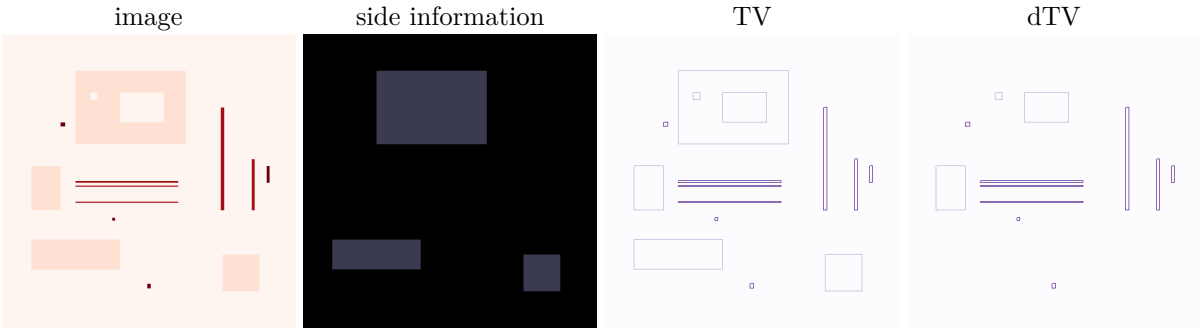

FIGURE 3.1. From left to right: *An image* $\boldsymbol{u}$*, side information* $\boldsymbol{v}$*, pointwise TV-norm* $j \mapsto \|(\nabla u)_j\|$*, and pointwise dTV-norm* $j \mapsto \|\mathbf{P}_j(\nabla \boldsymbol{u})_j\|$ *as in* (3.1).

3.2.1. Forward-backward splitting algorithm. The forward-backward splitting (FBS) algorithm [13] solves the composite minimization problem

$$\min_{\boldsymbol{u}} \{F(\boldsymbol{u}) + G(\boldsymbol{u})\} \tag{3.6}$$

where $F\colon X \to \mathbb{R}$ and $G\colon X \to (-\infty, \infty]$ are two proper, lower semi-continuous and convex functionals such that F is differentiable on X with a L-Lipschitz continuous gradient for some $L \in (0, \infty)$.

The principle of this algorithm is based on the two following steps:

1. a forward (explicit) gradient step on F, *i.e.* $\boldsymbol{u}^{t+1/2} = \boldsymbol{u}^t - \sigma^t \nabla F(\boldsymbol{u}^t)$, and
2. a backward (implicit) step involving only G, *i.e.* $\boldsymbol{u}^{t+1} = \operatorname{prox}_{\sigma^t G}(\boldsymbol{u}^{t+1/2})$, where the proximal operator is given by

$$\operatorname{prox}_{\sigma G}(\boldsymbol{z}) = \arg\min_{\boldsymbol{y}} \left\{ \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{z}\|^2 + \sigma G(\boldsymbol{y}) \right\}. \tag{3.7}$$

The step size σ^t is chosen in each iteration so that it satisfies the descent inequality

$$F\left(\boldsymbol{u}^{t+1}\right) \leq F\left(\boldsymbol{u}^t\right) + \left\langle \nabla F\left(\boldsymbol{u}^t\right), \boldsymbol{u}^{t+1} - \boldsymbol{u}^t \right\rangle + \frac{1}{2\sigma^t}\|\boldsymbol{u}^{t+1} - \boldsymbol{u}^t\|^2. \tag{3.8}$$

More precisely, we reduce σ^t until the condition (3.8) is satisfied. This selection of σ is known as *backtracking* and is considered in FBS and Bregman iterations.

Now, comparing problem (3.6) with (3.3), we choose the functions F and G as

$$F(\boldsymbol{u}) = \frac{1}{2}\|A\boldsymbol{u} - \boldsymbol{b}\|_2^2, \qquad G(\boldsymbol{u}) = \alpha J(\boldsymbol{u}) + \iota_{[0,\infty)^N}(\boldsymbol{u}).$$

We use the Fast Gradient Projection (FGP) algorithm presented in [3], to compute the proximal operator of the regularization functionals TV and dTV. For this latter, the detailed algorithm was presented in [17, Algorithm 1] where dTV was introduced for the first time. As described in [17], the minimization problem associated to the definition of proximal operator is dualized, this dual variable p is initialized as zero and following Algorithm 1 in [17], we set the number of (inner) iterations to be 200 and a tolerance of 10^{-5}, the first condition that is reached stops the algorithm. Additionally, we define the objective function value at point $\boldsymbol{u}$ as $H(\boldsymbol{u}) = F(\boldsymbol{u}) + G(\boldsymbol{u})$. Since FBS algorithm converges to a minimizer of H [13], we stop the algorithm when the difference between two consecutive iterations of the H value is less than a given tolerance tol, *i.e.* $H(\boldsymbol{u}^{t+1}) - H(\boldsymbol{u}^t) \leq \text{tol} \cdot H(\boldsymbol{u}^{t+1})$. The algorithm 1 describes one iteration of FBS algorithm.

Algorithm 1 An iteration of forward-backward splitting algorithm

$\boldsymbol{u}^{t+1} = \text{prox}_{\sigma^t G}\left(\boldsymbol{u}^t - \sigma^t \nabla F(\boldsymbol{u}^t)\right).$
if $F(\boldsymbol{u}^{t+1}) > F(\boldsymbol{u}^t) + \langle \nabla F(\boldsymbol{u}^t), \boldsymbol{u}^{t+1} - \boldsymbol{u}^t \rangle + \frac{1}{2\sigma^t}\|\boldsymbol{u}^{t+1} - \boldsymbol{u}^t\|^2$ **then**
    $\sigma^t = \underline{\rho}\sigma^t$, for any $\underline{\rho} < 1$ and go back to Step 2.
else
    $\sigma^{t+1} = \overline{\rho}\sigma^t$, for any $\overline{\rho} > 1$.
end if

3.3. Iterative regularization. A different way to achieve regularization is to apply an iterative method to directly solve the problem (2.1). Iterative methods start with a some vector $\boldsymbol{u}^0$ and generate a sequence $\boldsymbol{u}^1, \boldsymbol{u}^2, \ldots$ that converges to some solution. Usually in these methods, initial iterates $\boldsymbol{u}^t$ are fairly close to the exact solution. However, for later iterations, the solutions start to diverge from the desired one and tend to converge to the naive solution $A^{-1}\boldsymbol{b}$. Thus, the success of these methods relies on stopping the iterations at the right time. This behavior is known as *semiconvergence* [25, 41] and it is a frequently used tool to solve large-scale problems. In figure 3.2 we present an example of this effect. Additionally, iterative regularization avoids a predetermined regularization parameter, and instead, the number of iterations takes the role of a regularization parameter [25]. This is an advantage compared to variational regularization since in this latter, a minimization problem needs to be solved every time that a new regularization parameter α is tested.

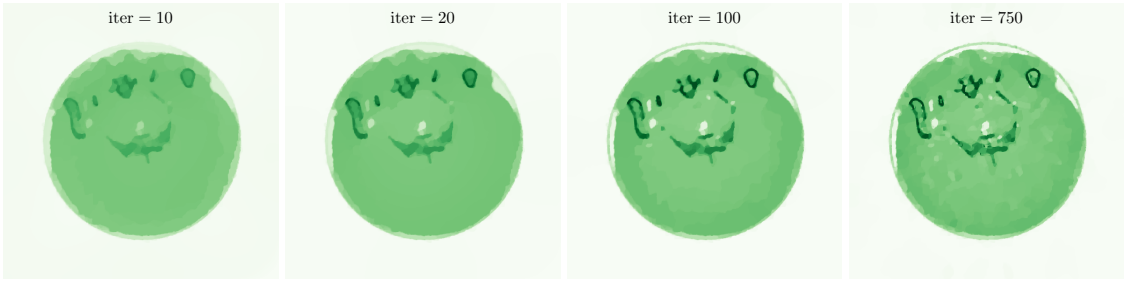

FIGURE 3.2. *Iterations along the Linearized Bregman iterations. While early iterations are very smooth, the iterates become gradually better defined and eventually the measurement noise is introduced.*

3.3.1. Linearized Bregman iterations. Under this group of iterative methods, we explore the Linearized Bregman iterations. This algorithm allows us to solve the least squares problem (2.1).

We consider the (linearized) Bregman iterations [43, 61] which makes use of the Bregman distance defined in terms of a given functional J by

$$D_J^{\boldsymbol{q}^t}(\boldsymbol{u}, \boldsymbol{u}^t) = J(\boldsymbol{u}) - J(\boldsymbol{u}^t) - \langle \boldsymbol{q}^t, \boldsymbol{u} - \boldsymbol{u}^t \rangle,$$

where $\boldsymbol{q}^t \in \partial J(\boldsymbol{u}^t)$ is an element of the sub-differential of J at point $\boldsymbol{u}^t$.

Linearized Bregman iterations are defined as

$$\boldsymbol{u}^{t+1} = \text{prox}_{\sigma^t G}\left(\boldsymbol{u}^t + \sigma^t(\boldsymbol{q}^t - \nabla F(\boldsymbol{u}^t))\right) \tag{3.9}$$

$$\boldsymbol{q}^{t+1} = \boldsymbol{q}^t - \frac{1}{\sigma^t}\left(\boldsymbol{u}^{t+1} - \boldsymbol{u}^t + \sigma^t \nabla F(\boldsymbol{u}^t)\right) \tag{3.10}$$

where $F(\boldsymbol{u}) = \frac{1}{2}\|A\boldsymbol{u} - \boldsymbol{b}\|_2^2$ is the objective function value and G can be chosen as G_{TV} or G_{dTV} from (3.4) and (3.5), respectively. The algorithm with backtracking is detailed in algorithm 2.

Algorithm 2 An iteration of Linearized Bregman iterations.

$\boldsymbol{u}^{t+1} = \operatorname{prox}_{\sigma^t G}(\boldsymbol{u}^t + \sigma^t(\boldsymbol{q}^t - \nabla F(\boldsymbol{u}^t)))$
if $F(\boldsymbol{u}^{t+1}) > F(\boldsymbol{u}^t) + \langle \nabla F(\boldsymbol{u}^t), \boldsymbol{u}^{t+1} - \boldsymbol{u}^t \rangle + \frac{1}{2\sigma^t}\|\boldsymbol{u}^{t+1} - \boldsymbol{u}^t\|^2$ **then**
    $\sigma^t = \underline{\rho}\sigma^t$, for any $\underline{\rho} < 1$ and go back to Step 2.
else
    $\sigma^{t+1} = \overline{\rho}\sigma^t$, for any $\overline{\rho} > 1$.
end if

4. Numerical results. We consider two sets of data, the first one, related to real measured data and a second one using synthetic (simulated) data. In both cases, we consider three energies labeled as E_0, E_1 and E_2, which are reconstructed separately. We analyze each energy channel independently as individual optimization problems. We compare the results from forward-backward splitting and linearized Bregman iterations and, highlight the main differences between TV and dTV regularizers. These algorithms were implemented using Python programming language and the Operator Discretization Library (ODL) [1]. For each energy channel, we consider sinograms of size $60 \times$ ndet, *i.e.* 60 projection angles and *ndet* detectors. The angles are uniformly distributed in the interval $[0, 2\pi)$ and the reconstructed images $\boldsymbol{u}$ are of size 512×512. For measured data, ndet $= 552$ and, for synthetic, ndet $= 640$.

First, we detail how to choose the side information $\boldsymbol{v}$ in our experiments considering the multispectral information in each energy channel.

4.0.1. Choice of side information. We propose to reconstruct a polyenergetic image $\boldsymbol{v} \in \mathbb{R}^N$ based on combining the data sets $\boldsymbol{b}_k \in \mathbb{R}^M$ for $k \in \{1, 2, 3\}$, *i.e.*, we solve

$$\boldsymbol{v} \in \operatorname*{arg\,min}_{\boldsymbol{u} \in \mathbb{R}^N} \left\{ \frac{1}{2}\|A\boldsymbol{u} - \tilde{\boldsymbol{b}}\|_2^2 + \alpha\,\mathrm{TV}(\boldsymbol{u}) + \iota_{[0,\infty)^N}(\boldsymbol{u}) \right\}. \tag{4.1}$$

where $\tilde{\boldsymbol{b}} = \sum_{k=1}^{3} \boldsymbol{b}_k$. The regularization parameter α and more details related to this optimization problem will be specified during the numerical experiments for synthetic and real data. Solving (4.1), we get an image $\boldsymbol{v}$ that despite of losing the spectral resolution, keeps structural information provided by all energy levels. Additionally, this image has higher signal-to-noise ratio and helps to improve the individual reconstructions $\boldsymbol{u}_k$ as we show in our experiments.

We present the results using red, green and blue color maps for E_0, E_1 and E_2, respectively. We use the color grey to distinguish everything related to side information, making an analogy with the grayscale representation of an RGB image.

4.1. Real data experiments. Experimental data was gathered at the Department of Physics, University of Helsinki, using a cone-beam micro-CT scanner with an end-window tube and a tungsten target (GE Phoenix nanotom 180 NF). The phantom used in this experiment corresponds to the cross section of a small bird (at chest level) with different tissue types and fine bone structures. This phantom was previously introduced in [55], it was imaged using three different X-ray tube settings averaging four frames in each projection to increase signal-to-noise ratio. The three energy datasets were generated using 50, 80 and 120 kV, respectively. For E_0, an electric current of 300 µA, an exposure time of 125 µs and no filtration were used; for E_1, the experiment was setted up with 180 µA, 125 µs and 1 mm of aluminium filtration. Similarly, for E_2, 120 µA, 250 µs and 0.5 mm of copper filtration were configured. 2D sinograms were created using the central plane of the cone-beam projections, in which the geometry reduces to a fan-beam geometry. First, we discuss about the choice of reference images and side information.

Reference images:. Given that the object of study was the cross section of a bird and not precisely a human patient, a complete scan was carried out for each energy, that is, an experiment with high radiation dose and high exposure time to observe 720 rotation angles. This provided us a high resolution image to be used as ground truth. More precisely, the reference image for each energy uses 720 angles and 552 detectors and is computed via (3.3) using FBS and TV regularizer. The regularization parameter α is chosen to preserve low noise and high-resolution details in the reconstructions. The proposed references are in figure 4.1. Alternative references can be obtained using Filtered-backprojection algorithm, in our case, since the choice of optimal α (for FBS) and optimal iteration (Bregman iterations) is related to the PSNR values, the use of a FBP reconstruction as reference did not give us good results, on the contrary, images with few details or a lot of noise were obtained.

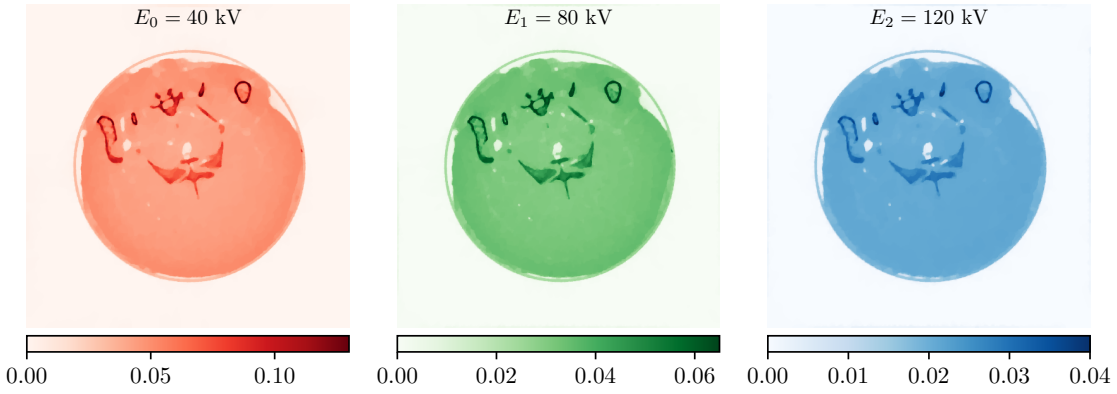

FIGURE 4.1. *Reference images for real data solving the problem* (3.3) *with* $\alpha = 0.005$, $\alpha = 0.002$ *and* $\alpha = 0.002$ *for* E_0, E_1 *and* E_2, *respectively.*

Choice of side information: . We solve the problem (4.1) for different values of α, we compare the resulting reconstructions in figure 4.2, also including a FBP-side information. We chose the reconstruction with $\alpha = 0.03$, that keeps sharper boundaries and includes few artifacts during the reconstruction. We will compare the potential differences when a FBP- and a TV-side informations are used in synthetic data section (see figure 4.14), we will mainly observe the improvements obtained by a regularized prior image.

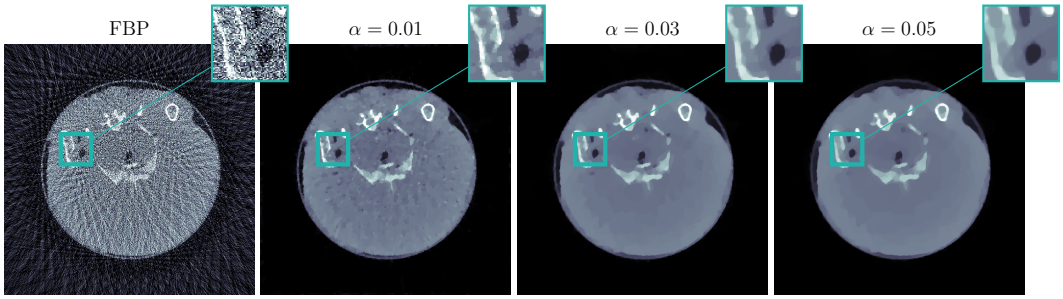

FIGURE 4.2. *From left to right: FBP reconstruction of side information with 60 angles and 552 detectors, TV solutions of the problem* (3.3) *with* $\alpha = 0.01$, $\alpha = 0.03$ *and* $\alpha = 0.05$, *respectively. The minimization problem is solved with FBS algorithm in a space of size* 512×512.

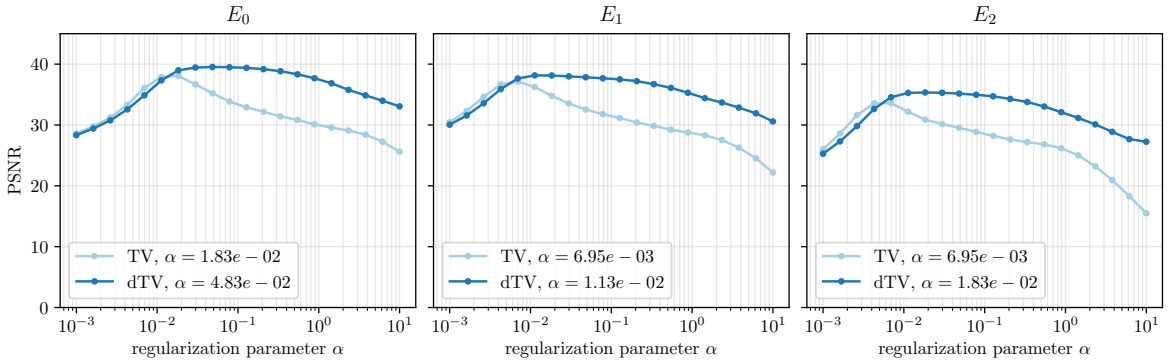

FIGURE 4.3. *PSNR values for 20 different regularization parameters using 60 angles. FBS algorithm is used with TV and dTV regularizers and the optimal α corresponds to the highest value of PSNR.*

FBS results:. We run the FBS iterations from algorithm 1, starting with $\boldsymbol{u}^0 = 1 \in \mathbb{R}^N$ and $\sigma^0 = 1/\|A\|^2$, where $\|A\|$ is an estimated norm of the operator A. We set tolerance as $\text{tol} = 10^{-6}$ for all the experiments, so the algorithm stops when $H(\boldsymbol{u}^{t+1}) - H(\boldsymbol{u}^t) \leq \text{tol} \cdot H(\boldsymbol{u}^{t+1})$. We choose $\eta = 0.01 \cdot \max_x |\nabla v(x)|$ for dTV definition (3.2) as commonly done for this regularizer [17, 7]. Since we are solving the problem (3.3), the parameter α is chosen by running an arrange of values between 10^{-3} and 10^1, the problem is solved for each value to compute the Peak Signal-to-Noise Ratio (PSNR) [27], as in figure 4.3. The optimal α corresponds to the highest PSNR. The results for all energies are shown in figure 4.4, we have included the structural similarity measure (SSIM) [56] and PSNR measures implemented in ODL, to compare the quality of the reconstructions to the reference images in figure 4.1.

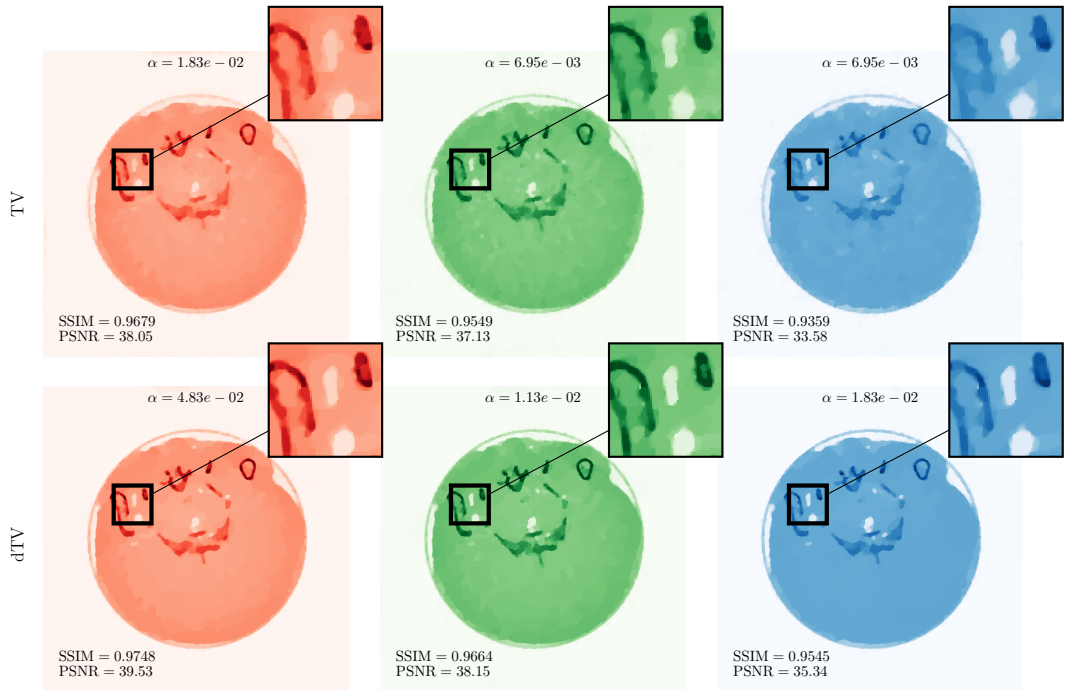

FIGURE 4.4. *For the three energies, reconstructions using FBS algorithm with TV (upper row) and dTV (bottom row). For each setting, α is chosen to maximize PSNR.*

Bregman results:. For algorithm 2, we start with $\sigma^0 = 1/\|A\|^2$ and $\boldsymbol{u}^0 = \boldsymbol{q}^0 = \mathbf{0} \in \mathbb{R}^N$. These choices guarantee that $\boldsymbol{q}^0 \in \partial G(\boldsymbol{u}^0)$ for $G = G_{\mathrm{TV}}$ or $G = G_{\mathrm{dTV}}$. We run 1000 iterations in algorithm 2 using $\alpha = 10$ for (3.4) and (3.5). We observe from figure 4.5 a common pattern along energies: the PSNR curve for TV is always below the curve associated to dTV, additionally, the number of iterations needed to maximize PSNR is always smaller for dTV than TV. The triangle markers refer to the best iterations in terms of PSNR. The images with highest PSNR are presented in figure 4.6.

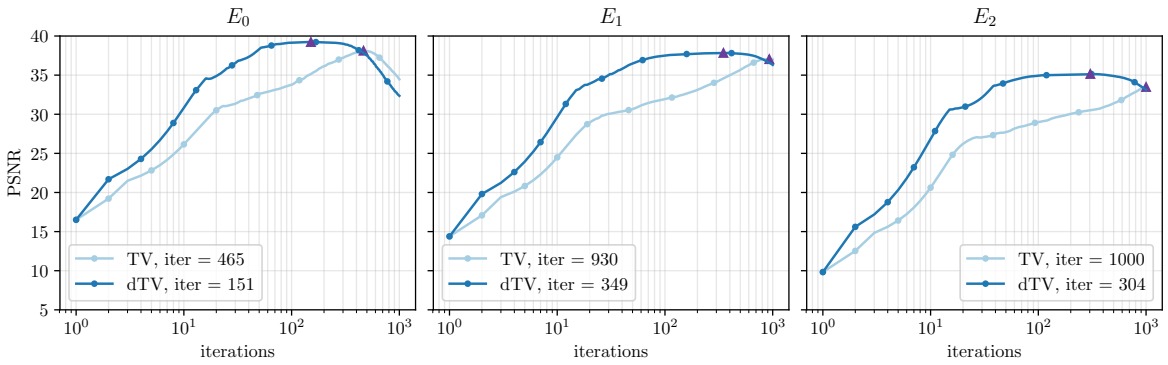

FIGURE 4.5. *For each energy level, the graphs show iterations against PSNR.*

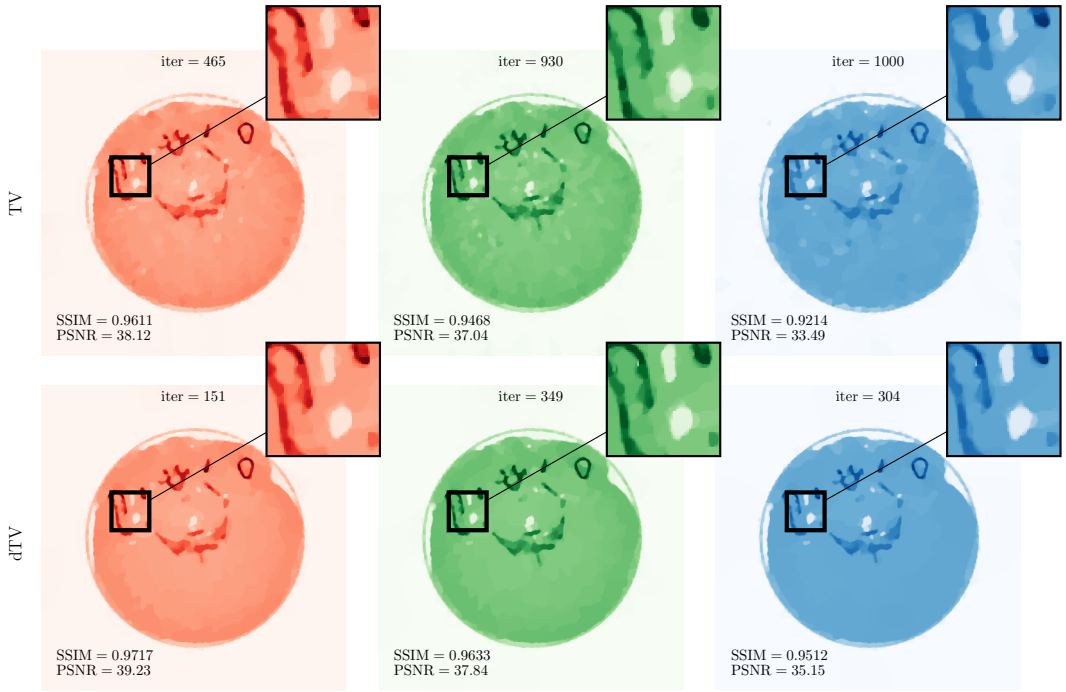

FIGURE 4.6. *For the three energies, reconstructions using Bregman iterations with TV(upper row) and dTV (bottom row). Each image is labeled by the iteration that maximize PSNR.*

Comparison between FBS and Bregman iterations:. After calibration of the regularization parameter α for FBS and iteration number for Bregman iterations, we compare the two algorithms for E_2 in figure 4.7. We observe that both algorithms give similar values for PSNR and SSIM but consistently dTV outperforms TV.

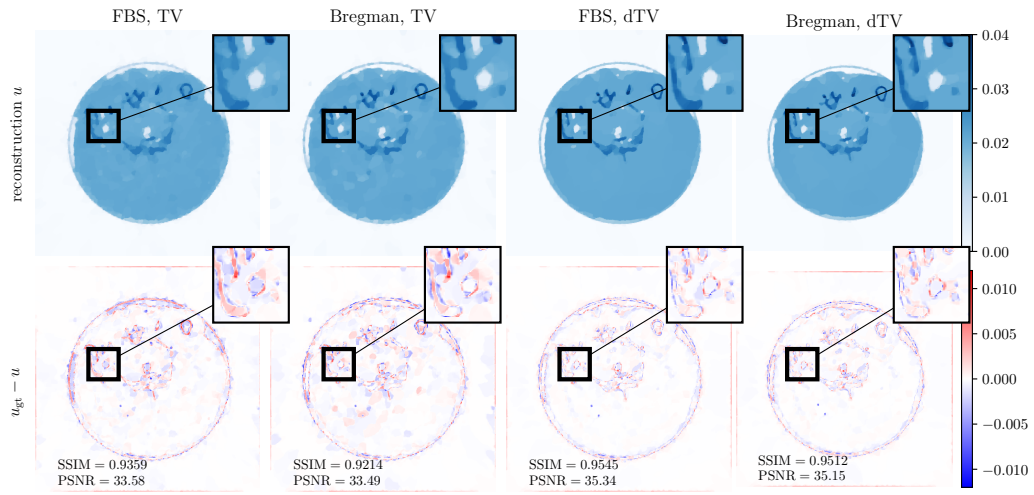

FIGURE 4.7. *Top: reconstructions for E_2 using "optimal" regularization for both FBS and Bregman iterations. Bottom: difference between reconstruction and reference image (see figure 4.1).*

4.2. Synthetic data experiments. For the simulated data, we used an anthropomorphic chest phantom generated using the XCAT software [50]. We simulated measurements using the XCAT CT Simulator, using an X-ray spectrum included with the software. The simulated spectrum was created using SRS-78 Spectrum Processor program, and it modelled a tungsten tube X-ray source with a 120 kV acceleration voltage and 5 mm of aluminum filtration. For the multi-energy measurement simulations, the spectrum was divided into three energy bins: 0-60 kV, 60- 90 kV, and 90-120 kV. For testing the algorithms, we created 720 projections at 0.5 degree intervals, with simulated photon noise. For creating the reference/ground truth, we simulated 1440 projections at 0.25 degree intervals, with no noise. In figure 4.8, we present the energy spectrum for our data and in figure 4.9 the reference images.

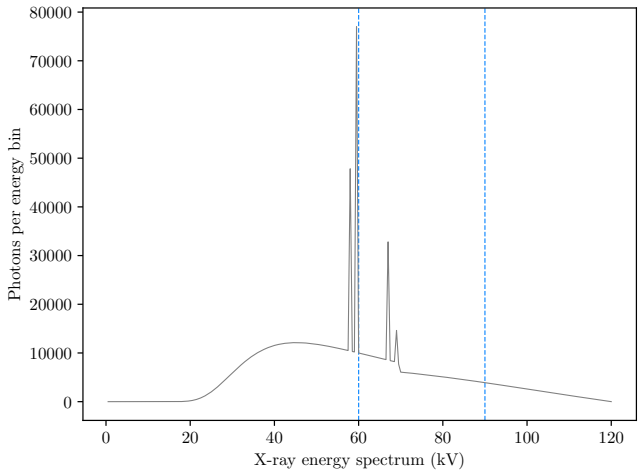

FIGURE 4.8. *The energy spectrum with the three energy bins divided at* 60 kV *and* 90 kV.

Choice of side information: . As in real case, we solve the problem (4.1) using $\alpha = 0.08$, $\alpha = 0.1$ and $\alpha = 0.5$. We have chosen the image with $\alpha = 0.1$, which gave us the best results. We discuss the accuracy of side information choice in figure 4.14.

FBS results:. We initialized $\boldsymbol{u}^0$ and σ^0 as in real experiments. In figure 4.10, we present the PSNR values obtained for 20 different values of α from 10^{-3} to 10^2 for TV and dTV regularizers in each energy channel. The corresponding reconstructions are in figure 4.11. We included close-ups for easier comparison of the reconstructions. Here, we observe that using dTV yields higher values of measures and better reconstructions reducing the artifacts in the zoomed area at the top.

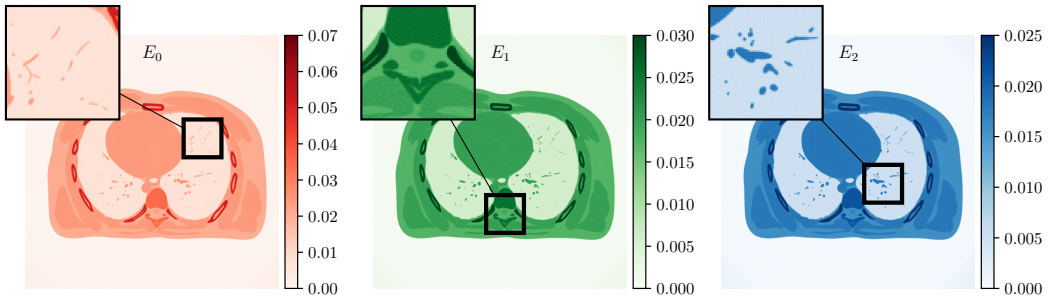

FIGURE 4.9. From left to right: *The energy spectrum with the three energy bins divided at* 60 kV *and* 90 kV *and, reference images for* E_0, E_1 *and* E_2. *Each image presents a different zoomed zone as a reference for the oncoming experiments.*

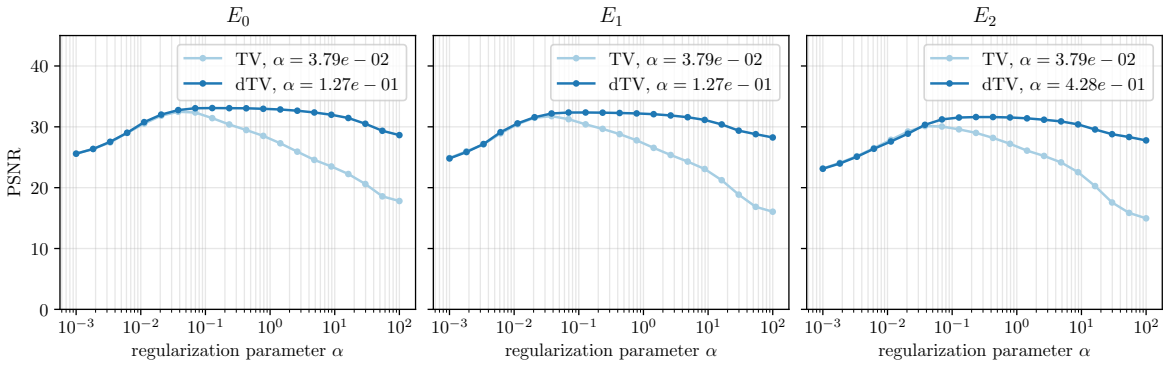

FIGURE 4.10. *PSNR values for 20 different regularization parameters using 60 angles. FBS algorithm is used with TV and dTV regularizers and the optimal* α *corresponds to the highest value of PSNR.*

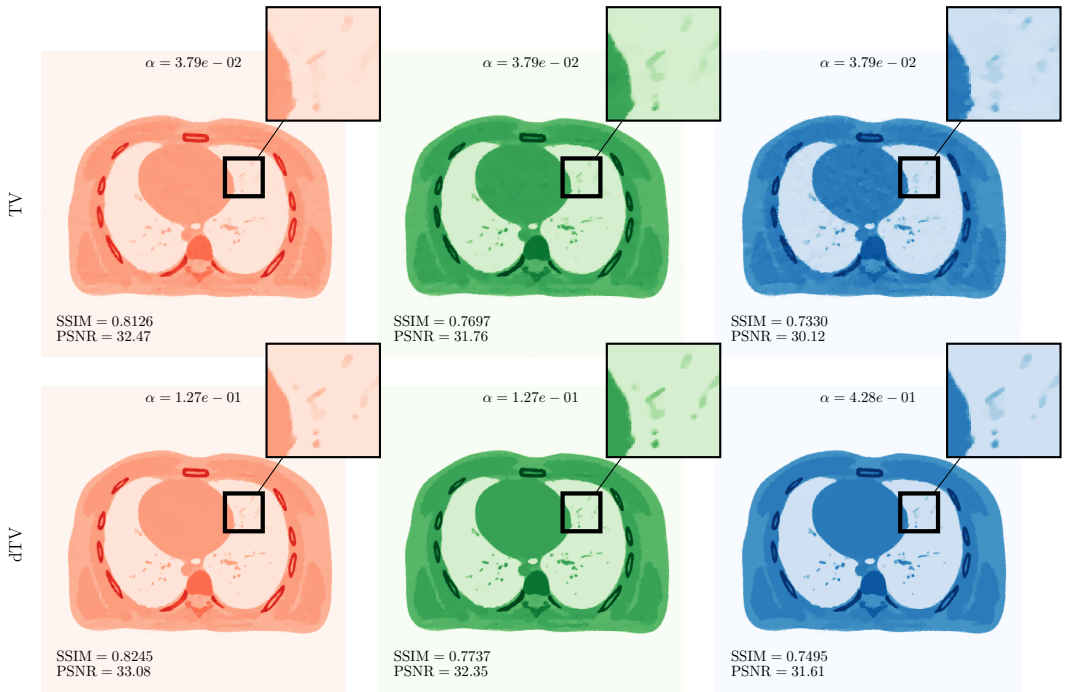

FIGURE 4.11. *For the three energies, reconstructions using Bregman iterations with TV (upper row) and dTV (bottom row). Each image is labeled by the iteration that maximize PSNR.*

Bregman results:. Now, we observe the results for algorithm 2. We initialize $\sigma^0 = 1/\|A\|^2$ and $\boldsymbol{u}^0 = \boldsymbol{q}^0 = \mathbf{0} \in \mathbb{R}^N$. We run the algorithm for 1000 iterations using three different regularization parameters, $\alpha = 1$, $\alpha = 10$, and $\alpha = 100$. As before, the "optimal" iteration number is chosen to maximize PSNR. In figure 4.12 (left), using E_1 as an example, we present the PSNR values along iterations for TV and dTV when these three values of α are considered in the minimization functional. We observe the influence of α in terms of number of iterations and quality of the reconstruction. While choosing a small value of α allows us to reach a maximum value of PSNR in a small number of iterations, the PSNR values are smaller compared to a higher α value. The highest values of PSNR occur when $\alpha = 100$ but TV reconstructions need more than 1000 of iterations to find a maximum PSNR. In figure 4.13, we present the resulting reconstructions for TV and dTV in E_1. We have chosen $\alpha = 10$ since the PSNR values in this case are close to the ones obtained with $\alpha = 100$ and the number of iterations are less than 250.

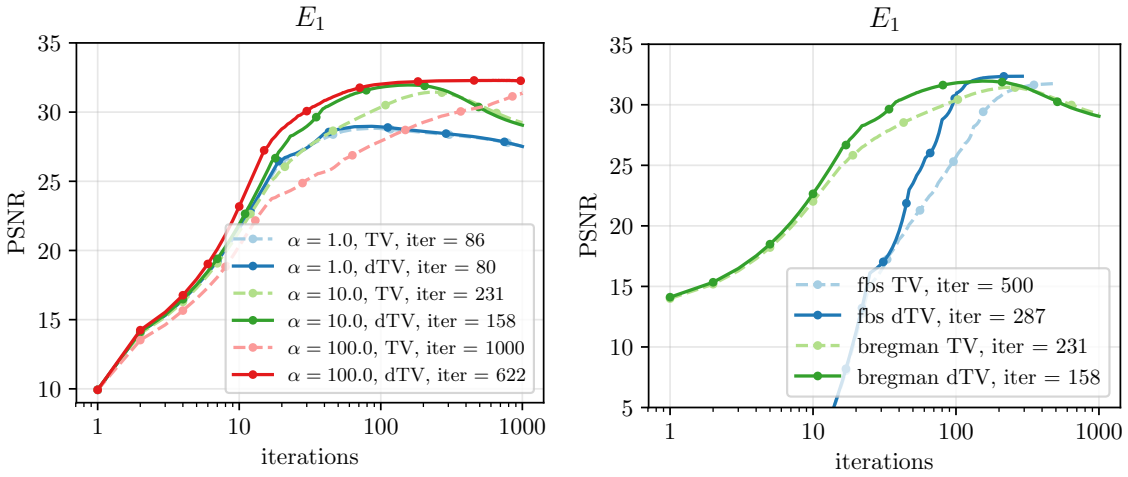

FIGURE 4.12. Left: *PSNR values for Bregman iterations using $\alpha = 1$, $\alpha = 10$, and $\alpha = 100$. The results are shown for both regularizers and the optimal iteration is included based on the highest PSNR value.* Right: *PSNR along iterations for FBS and Bregman iterations using both TV and dTV.*

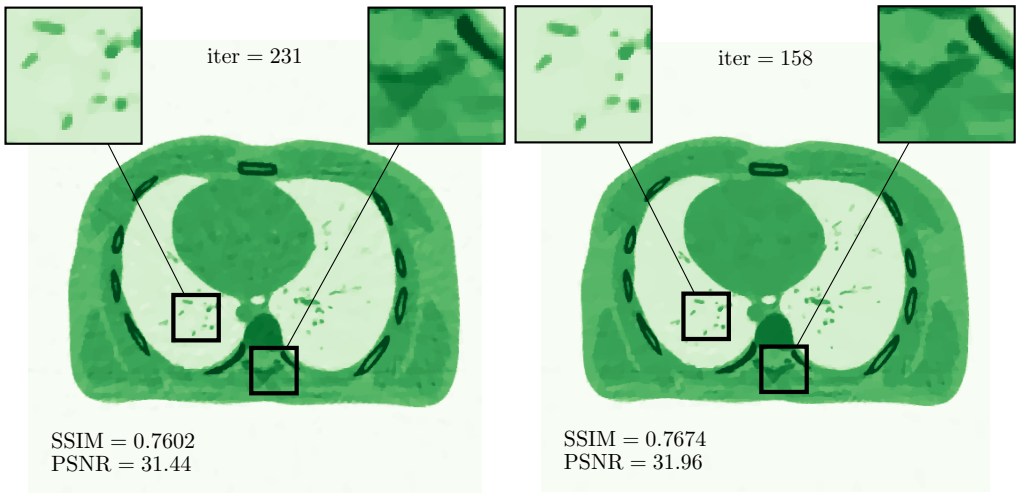

FIGURE 4.13. *Bregman iterations for TV (left) and dTV (right) for E_1 using $\alpha = 10$ based on the left hand of figure 4.12.*

The FBS and Bregman reconstructions for E_1 in figure 4.11 and figure 4.13 reach similar results, however, we note that fewer iterations were needed for Bregman iterations compared to FBS as shown in figure 4.12 (right). This means that linearized Bregman iterations converge faster to desired solution than forward-backward splitting.

4.2.1. Influence of side information. In this experiment, we compare the accuracy of the reconstructions depending on the choice of side information. For this, we consider a FBP- an TV-side information reconstructions. The second one was obtained by solving (4.1) with $\alpha = 0.1$. These two choices, shown in figure 4.14 for XCAT phantom, give us two types of side information, one image with more structures and artifacts (upper row) and another image with smoother shapes but without artifacts (bottom row). We compare the best reconstructions using dTV with FBS algorithm considering the highest values of PSNR. For the first side information, the artifacts remain in the reconstructed image and poor quality values of PSNR are reached using larger number of Bregman iterations compared to the second case.

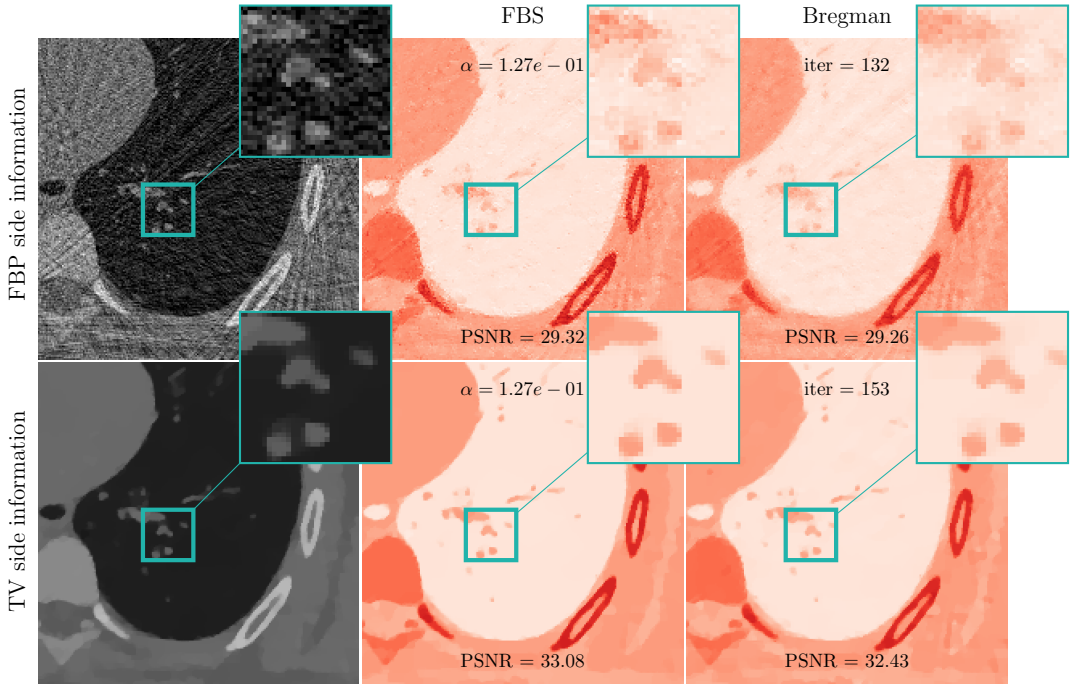

FIGURE 4.14. *Comparison between dTV reconstructions using FBS and Bregman iterations with two different side informations, one TV-regularized with $\alpha = 100$ and a second, non-regularized side information with FBP.*

5. Conclusion. We have analyzed synergistic reconstruction for multi-spectral CT reconstruction when a limited set of angles is observed. The proposed approach is based on combining information from all available energy channels into a polyenergetic image. This image is then included into the directional total variation regularizer for use in variational or iterative regularization.

We observed that the synergistic approach based on directional total variation is always superior to separate reconstruction using just total variation for both variational and iterative regularization. In addition, we consistently saw that linearized Bregman iterations converge faster to a desired solution than forward-backward splitting.

Additionally, reconstructing images channel-by-channel due to dTV with channel-independent side information allows us to access the computational advantages of parallel programming, differentiating our proposal from methods that solve the multi-spectral problem by a single large optimization problem.

The observation that synergistic reconstruction can be faster than separate reconstruction is novel and interesting and future work will be directed to fully understand this phenomenon.

Appendix: Comparison with TVN. Comparing the results with other methodologies used for multi-spectral CT is pertinent, for example with total nuclear variation (TNV) described in [45] on realistic dual-energy CT data. To do this, we solved the problem:

$$\boldsymbol{u}^* \in \operatorname*{arg\,min}_{\boldsymbol{u}\in(\mathbb{R}^N)^3} \left\{ \sum_{i=0}^{2} \|R\boldsymbol{u}_i - \boldsymbol{b}_i\|_2^2 + \alpha \mathrm{TNV}(\boldsymbol{u}) + \iota_{[0,\infty)^N}(\boldsymbol{u}) \right\}.$$

where $\boldsymbol{u} = (\boldsymbol{u}_0, \boldsymbol{u}_1, \boldsymbol{u}_2)$ correspond to jointly find the reconstruction for the three images for E_0, E_1 and E_2. The results obtained are presented and compared with TV and dTV in Figure 5.1. The optimization problem was solved using Primal-dual hybrid gradient (pdhg) algorithm described in [9]. More details related to TNV with pdhg can be found in [15].

Comparing PSNR values for TNV and dTV, we can observe that the differences are quite negligible. In the case of dTV, sharper edges can be observed compared to TV and TNV. We could conclude that TNV and dTV methods give similar results, however, there are important differences in terms of computational cost. In TVN a large problem needs to be solved, if we have to reconstruct N energy channels all have to be solve in the same computing unit, while using dTV in a channel-by-channel reconstruction allows us to solve these N problems in N different computers (nodes) through a parallelization process.

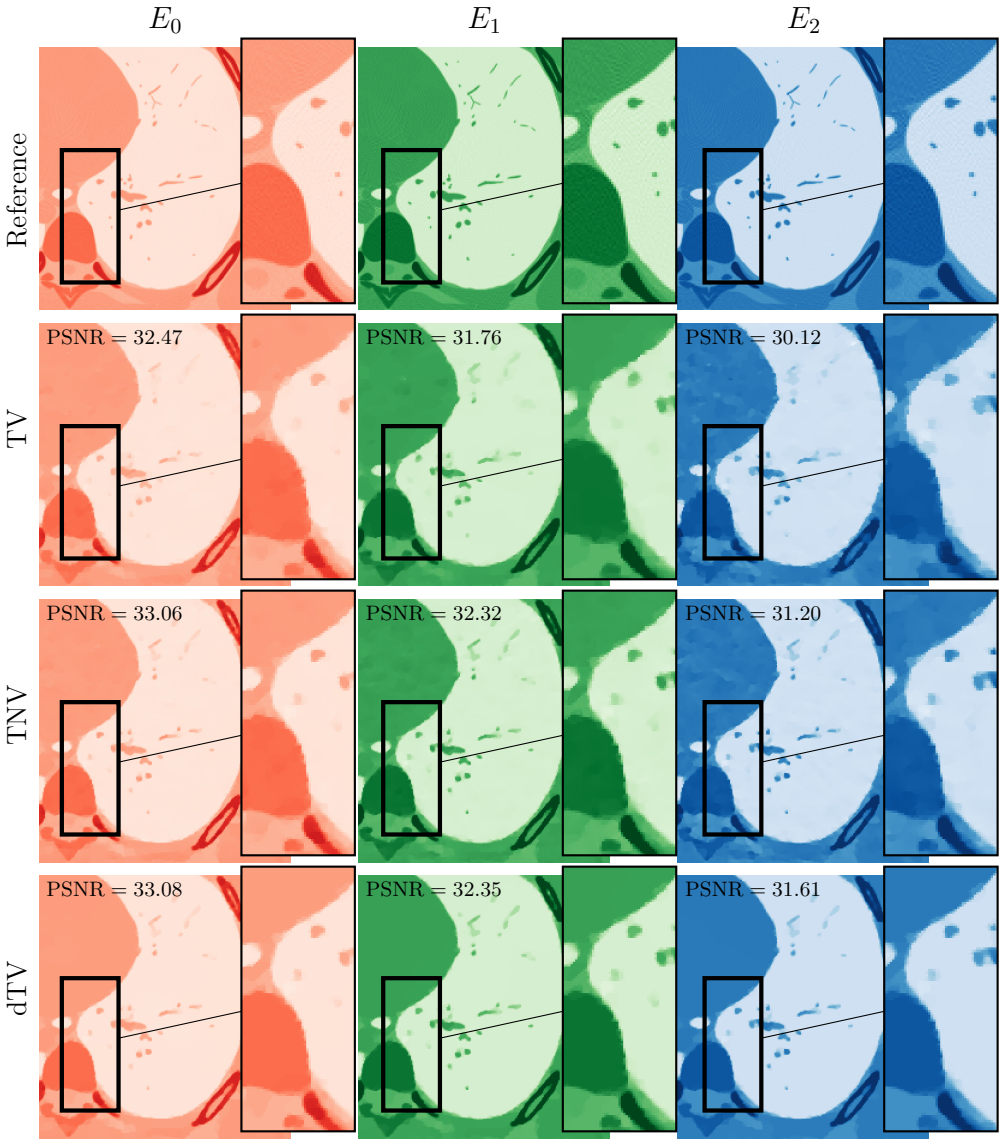

FIGURE 5.1. *Comparison between TV, TVN and dTV regularizers. The PSNR values are included in each reconstruction. The noiseless images (references) for each energy are included in the top row.*

Authors' Contributions: EC carried out the numerical computations and drafted the manuscript. MJE designed the project, advised EC and supported writing the manuscript. SS and AM acquiered and processed real data. All authors read and approved the manuscript.

Funding: MJE acknowledges support from the EPSRC (EP/S026045/1, EP/T026693/1), the Faraday Institution (EP/T007745/1) and the Leverhulme Trust (ECF-2019-478). EC acknowledges support from the CMM ANID PIA AFB170001 and Beca Doctorado Nacional Conicyt.

REFERENCES

[1] J. Adler, H. Kohr, and O. Oktem. Operator discretization library (odl). *Software available from https://github.com/odlgroup/odl*, 2017. URL `https://github.com/odlgroup/odl`.

[2] A. M. Alessio and L. R. MacDonald. Quantitative material characterization from multi-energy photon counting ct. *Medical physics*, 40(3):031108, 2013. doi:10.1118/1.4790692.

[3] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm with application to wavelet-based image deblurring. In *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 693–696. IEEE, IEEE, apr 2009a. doi:10.1109/icassp.2009.4959678.

[4] M. Benning and M. Burger. Ground states and singular vectors of convex variational regularization methods. *Methods and Applications of Analysis*, 20(4):295–334, 2013. doi:10.4310/maa.2013.v20.n4.a1.

[5] M. Benning, M. M. Betcke, M. J. Ehrhardt, and C.-B. Schönlieb. Choose your path wisely: gradient descent in a Bregman distance framework. *arXiv preprint arXiv:1712.04045*, 2017.

[6] T. Bonesky. Morozov's discrepancy principle and tikhonov-type functionals. *Inverse Problems*, 25(1):015015, dec 2008. doi:10.1088/0266-5611/25/1/015015.

[7] L. Bungert, D. A. Coomes, M. J. Ehrhardt, J. Rasch, R. Reisenhofer, and C.-B. Schönlieb. Blind image fusion for hyperspectral imaging with the directional total variation. *Inverse Problems*, 34(4):044003, 2018. doi:10.1088/1361-6420/aaaf63.

[8] L. D. Chiffre, S. Carmignato, J.-P. Kruth, R. Schmitt, and A. Weckenmann. Industrial applications of computed tomography. *CIRP Annals*, 63(2):655–677, 2014. doi:10.1016/j.cirp.2014.05.011.

[9] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of mathematical imaging and vision* 40(1):120–145, 2011. doi:10.1007/s10851-010-0251-1.

[10] R. Clackdoyle and M. Defrise. Tomographic reconstruction in the 21st century. *IEEE Signal Processing Magazine*, 27(4):60–80, 2010. doi:10.1109/msp.2010.936743.

[11] D. Colton, M. Piana, and R. Potthast. A simple method using morozov's discrepancy principle for solving inverse scattering problems. *Inverse Problems*, 13(6):1477, 1997. doi:10.1088/0266-5611/13/6/005.

[12] P. L. Combettes and J.-C. Pesquet. Proximal splitting methods in signal processing. In *Springer Optimization and Its Applications*, pages 185–212. Springer New York, 2011. doi:10.1007/978-1-4419-9569-8_10.

[13] P. L. Combettes and V. R. Wajs. Signal recovery by proximal forward-backward splitting. *Multiscale Modeling & Simulation*, 4(4):1168–1200, 2005. doi:10.1137/050626090.

[14] V. Corona, M. Benning, M. J. Ehrhardt, L. F. Gladden, R. Mair, A. Reci, A. J. Sederman, S. Reichelt, and C.-B. Schönlieb. Enhancing joint reconstruction and segmentation with non-convex Bregman iteration. *Inverse Problems*, 35(5):055001, 2019a. doi:10.1088/1361-6420/ab0b77.

[15] M. J. Ehrhardt. Multi-modality imaging with structure-promoting regularisers. pages 36–43, 2020. URL `http://arxiv.org/abs/2007.11689`.

[16] M. J. Ehrhardt and S. R. Arridge. Vector-valued image processing by parallel level sets. *IEEE Transactions on Image Processing*, 23(1):9–18, jan 2014. doi:10.1109/tip.2013.2277775.

[17] M. J. Ehrhardt and M. M. Betcke. Multicontrast mri reconstruction with structure-guided total variation. *SIAM Journal on Imaging Sciences*, 9(3):1084–1106, 2016. doi:10.1137/15m1047325.

[18] M. J. Ehrhardt, K. Thielemans, L. Pizarro, D. Atkinson, S. Ourselin, B. F. Hutton, and S. R. Arridge. Joint reconstruction of pet-mri by exploiting structural similarity. *Inverse Problems*, 31(1):015001, 2014. doi:10.1088/0266-5611/31/1/015001.

[19] M. J. Ehrhardt, P. Markiewicz, M. Liljeroth, A. Barnes, V. Kolehmainen, J. S. Duncan, L. Pizarro, D. Atkinson, B. F. Hutton, S. Ourselin, K. Thielemans, and S. R. Arridge. PET reconstruction with an anatomical MRI prior using parallel level sets. *IEEE Transactions on Medical Imaging*, 35(9):2189–2199, sep 2016. doi:10.1109/tmi.2016.2549601.

[20] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems*, volume 375. Springer Netherlands, 1996. doi:10.1007/978-94-009-1740-8.

[21] J. Frikel and E. T. Quinto. Characterization and reduction of artifacts in limited angle tomography. *Inverse Problems*, 29(12):125007, 2013. doi:10.1088/0266-5611/29/12/125007.

[22] H. Gao, H. Yu, S. Osher, and G. Wang. Multi-energy CT based on a prior rank, intensity and sparsity model (PRISM). *Inverse Problems*, 27(11):115012, oct 2011. doi:10.1088/0266-5611/27/11/115012.

[23] C. Groetsch. Comments on morozov's discrepancy principle. In *Improperly posed problems and their numerical treatment*, pages 97–104. Springer, 1983. doi:10.1007/978-3-0348-5460-3_7.

[24] K. Hamalainen, A. Kallonen, V. Kolehmainen, M. Lassas, K. Niinimaki, and S. Siltanen. Sparse tomography. *SIAM Journal on Scientific Computing*, 35(3):B644–B665, 2013. doi:10.1137/120876277.

[25] P. C. Hansen. *Discrete Inverse Problems*, volume 7. Society for Industrial and Applied Mathematics, jan 2010. doi:10.1137/1.9780898718836.

[26] M. Hintermüller, M. Holler, and K. Papafitsoros. A function space framework for structural total variation regularization with applications in inverse problems. *Inverse Problems*, 34(6):064002, apr 2018. doi:10.1088/1361-6420/aab586.

[27] A. Horé and D. Ziou. Image Quality Metrics: PSNR vs. SSIM. *2020 20th International Conference on Pattern Recognition*, Istanbul, 2010, pp. 2366—2369, doi:10.1109/ICPR.2010.579.

[28] Y. Hu, J. G. Nagy, J. Zhang, and M. S. Andersen. Nonlinear optimization for mixed attenuation polyenergetic image reconstruction. *Inverse Problems*, 35(6):064004, jun 2019. doi:10.1088/1361-6420/ab0131.

[29] R. H. Huesman. The effects of a finite number of projection angles and finite lateral sampling of projections on the propagation of statistical errors in transverse section reconstruction. *Physics in Medicine and Biology*, 22 (3):511–521, may 1977. doi:10.1088/0031-9155/22/3/012.

[30] T. R. Johnson. Dual-energy ct: general principles. *American Journal of Roentgenology*, 199(5_supplement):S3–S8, 2012. doi:10.2214/ajr.12.9116.

[31] J. P. Kaipio, V. Kolehmainen, M. Vauhkonen, and E. Somersalo. Inverse problems with structural prior information. *Inverse Problems*, 15(4):1111–1111, aug 1999. doi:10.1088/0266-5611/15/4/501.

[32] D. Kazantsev, J. S. Jørgensen, M. S. Andersen, W. R. Lionheart, P. D. Lee, and P. J. Withers. Joint image reconstruction method with correlative multi-channel prior for X-ray spectral computed tomography. *Inverse Problems*, 34(6):064001, 2018. doi:10.1088/1361-6420/aaba86.

[33] B. Krauss, B. Schmidt, and T. G. Flohr. Dual source CT. In *Dual Energy CT in Clinical Practice*, pages 11–20. Springer Berlin Heidelberg, sep 2010. doi:10.1007/174_2010_44.

[34] D. Lee, J. Lee, H. Kim, T. Lee, J. Soh, M. Park, C. Kim, Y. J. Lee, and S. Cho. A feasibility study of low-dose single-scan dual-energy cone-beam CT in many-view under-sampling framework. *IEEE Transactions on Medical Imaging*, 36(12):2578–2587, dec 2017. doi:10.1109/tmi.2017.2765760.

[35] M. M. Lell, J. E. Wildberger, H. Alkadhi, J. Damilakis, and M. Kachelriess. Evolution in computed tomography. *Investigative Radiology*, 50(9):629–644, sep 2015. doi:10.1097/rli.0000000000000172.

[36] A. Macovski, R. Alvarez, J.-H. Chan, J. Stonestrom, and L. Zatz. Energy dependent reconstruction in X-ray computerized tomography. *Computers in biology and medicine*, 6(4):325–336, 1976. doi:10.1016/0010-4825(76)90069-x.

[37] D. Marin, D. T. Boll, A. Mileto, and R. C. Nelson. State of the art: Dual-energy CT of the abdomen. *Radiology*, 271(2):327–342, may 2014. doi:10.1148/radiol.14131480.

[38] C. H. McCollough, G. H. Chen, W. Kalender, S. Leng, E. Samei, K. Taguchi, G. Wang, L. Yu, and R. I. Pettigrew. Achieving routine submillisievert ct scanning: report from the summit on management of radiation dose in ct. *Radiology*, 264(2):567–580, 2012. doi:10.1148/radiol.12112265.

[39] C. H. McCollough, S. Leng, L. Yu, and J. G. Fletcher. Dual- and multi-energy CT: Principles, technical approaches, and clinical applications. *Radiology*, 276(3):637–653, sep 2015. doi:10.1148/radiol.2015142631.

[40] V. A. Morozov. *Methods for solving incorrectly posed problems*. Springer Science & Business Media, 2012.

[41] F. Natterer. *The Mathematics of Computerized Tomography*. Society for Industrial and Applied Mathematics, jan 2001. doi:10.1137/1.9780898719284.

[42] S. Niu, G. Yu, J. Ma, and J. Wang. Nonlocal low-rank and sparse matrix decomposition for spectral CT reconstruction. *Inverse Problems*, 34(2):024003, jan 2018. doi:10.1088/1361-6420/aa942c.

[43] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin. An iterative regularization method for total variation-based image restoration. *Multiscale Modeling & Simulation*, 4(2):460–489, jan 2005b. doi:10.1137/040605412.

[44] D. S. Rigie and P. J. L. Rivière. Joint reconstruction of multi-channel, spectral CT data via constrained total nuclear variation minimization. *Physics in Medicine and Biology*, 60(5):1741–1762, feb 2015. doi:10.1088/0031-9155/60/5/1741.

[45] D. S. Rigie, A. A. Sanchez and P. J. L. Rivière. Assessment of vectorial total variation penalties on realistic dual-energy CT data. *MPhysics in Medicine and Biology*, 62(8):3284–3298. doi:10.1088/1361-6560/aa6392.

[46] J. Rogowska. *Handbook of Medical Imaging: Processing and Analysis Management*, volume 5. 2000.

[47] L. I. Rudin and S. Osher and E. Fatemi. Nonlinear total variation based noise removal algorithms *Physica D: Nonlinear Phenomena*, Elsevier BV, 60, 259–268, 1992. doi:10.1016/0167-2789(92)90242-f.

[48] T. Sarkar, D. Weiner, and V. Jain. Some mathematical considerations in dealing with the inverse problem. *IEEE Transactions on Antennas and Propagation*, 29(2):373–379, mar 1981. doi:10.1109/tap.1981.1142573.

[49] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen. *Variational methods in imaging*. Springer, 2009.

[50] W. P. Segars, M. Mahesh, T. J. Beck, E. C. Frey, B. M. Tsui. Realistic CT simulation using the 4D XCAT phantom. *Med Phys.*, 35(8):3800–8, aug 2008. doi:10.1118/1.2955743.

[51] O. Semerci, N. Hao, M. E. Kilmer, and E. L. Miller. Tensor-based formulation and nuclear norm regularization for multienergy computed tomography. *IEEE Transactions on Image Processing*, 23(4):1678–1693, apr 2014. doi:10.1109/tip.2014.2305840.

[52] P. M. Shikhaliev. Energy-resolved computed tomography: first experimental results. *Physics in Medicine and Biology*, 53(20):5595–5613, sep 2008. doi:10.1088/0031-9155/53/20/002.

[53] P. M. Shikhaliev and S. G. Fritz. Photon counting spectral CT versus conventional CT: comparative evaluation for breast imaging application. *Physics in Medicine and Biology*, 56(7):1905–1930, mar 2011. doi:10.1088/0031-9155/56/7/001.

[54] D. Strong and T. Chan. Edge-preserving and scale-dependent properties of total variation regularization. *Inverse Problems*, 19(6):S165–S187, nov 2003. doi:10.1088/0266-5611/19/6/059.

[55] J. Toivanen, , A. Meaney, S. Siltanen, and V. K. and. Joint reconstruction in low dose multi-energy CT. *Inverse Problems & Imaging*, 14(4):607–629, 2020. doi:10.3934/ipi.2020028.

[56] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, April 2004, doi:10.1109/TIP.2003.819861.

[57] Y.-W. Wen and R. H. Chan. Parameter selection for total-variation-based image restoration using discrepancy principle. *IEEE Transactions on Image Processing*, 21(4):1770–1781, apr 2012. doi:10.1109/tip.2011.2181401.

[58] W. Wu, Y. Zhang, Q. Wang, F. Liu, F. Luo, and H. Yu. Spatial-spectral cube matching frame for spectral CT reconstruction. *Inverse Problems*, 34(10):104003, aug 2018. doi:10.1088/1361-6420/aad67b.

[59] D. Xu, D. A. Langan, X. Wu, J. D. Pack, T. M. Benson, J. E. Tkaczky, and A. M. Schmitz. Dual energy ct via fast kvp switching spectrum estimation. In *Medical Imaging 2009: Physics of Medical Imaging*, volume 7258, page 72583T. International Society for Optics and Photonics, 2009. doi:10.1117/12.811650.

[60] W. Yin, S. Osher, D. Goldfarb, and J. Darbon. Bregman iterative algorithms for ℓ_1-minimization with applications to compressed sensing. *SIAM Journal on Imaging Sciences*, 1(1):143–168, jan 2008. doi:10.1137/070703983.

[61] W. Yin. Analysis and generalizations of the linearized Bregman method. *SIAM Journal on Imaging Sciences*, 3 (4):856–877, 2010. doi:10.1137/090760350.

[62] Y. Zou and M. D. Silver. Analysis of fast kV-switching in dual energy CT using a pre-reconstruction decomposition technique. In J. Hsieh and E. Samei, editors, *Medical Imaging 2008: Physics of Medical Imaging*, volume 6913, page 691313. International Society for Optics and Photonics, SPIE, mar 2008. doi:10.1117/12.772826.